

# Webinar: Edición Masiva de Metadatos con OpenRefine



06/12/2019

# Sumário

**1 oasisbr y BDTD**

**2 DSpace en Brasil y Curación por medio del  
OpenRefine**

# oasisbr y BDTD - Visión general

- Presentación de los **portales de búsqueda**
- Composición de la **red**
- *Background* tecnologico y **interoperabilidad**
- Numero de **usuarios**

## oasisbr y BDTD

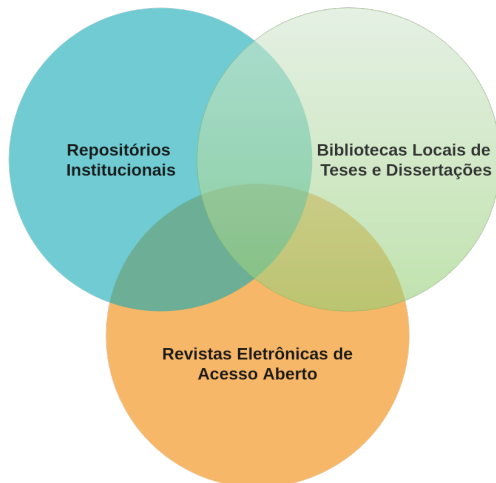
- Portal Brasileño de Publicaciones Científicas de Acceso Abierto (**oasisbr**)



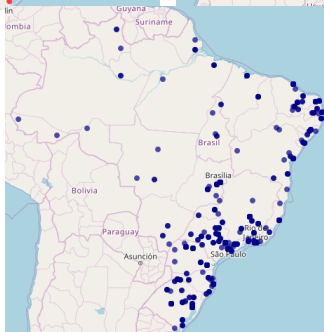
- Biblioteca Digital Brasileira de Teses y Disertaciones (**BDTD**)



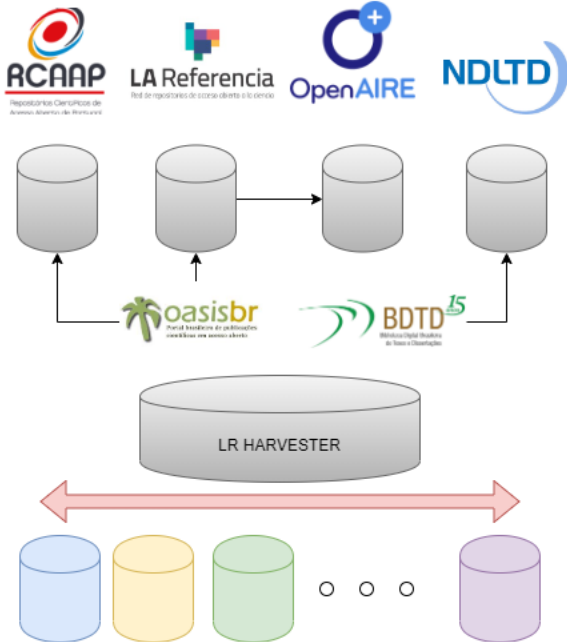
Red con más de **500** instituciones que agregan más de **1,250** fuentes y aproximadamente **2,300,000** docs.



	<b>RI</b> s	<b>TD</b> s	<i>Revistas</i>	<b>RCAAP</b>
<b># fontes</b>	<b>100</b>	<b>115</b>	<b>1,030</b>	<b>1</b>
<b># docs.</b>	<b>605K</b>	<b>580K</b>	<b>520K</b>	<b>595K</b>



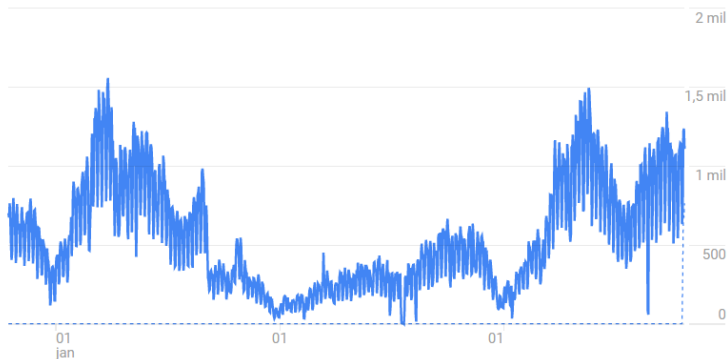
- Bibliotecas locais de TDs
- Repositórios Institucionais
- Revistas Eletrônicas





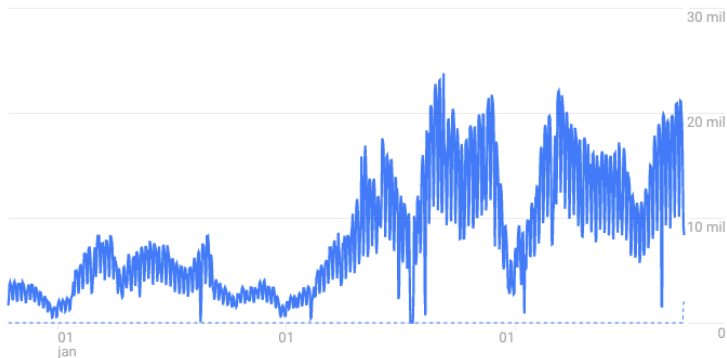
# Número de usuarios de oasisbr

out/2016 a out/2019



# Número de usuarios de BDTD

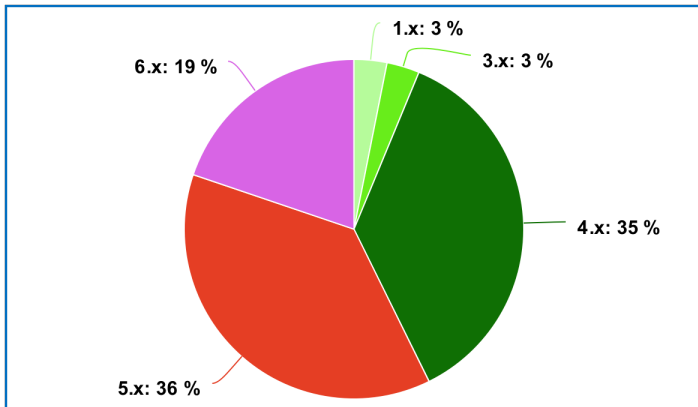
out/2016 a out/2019



# DSpace no Brasil – Visión general

- **Versiones** existentes
- **Distribución** geográfica
- **TEDE** y **DSpace *Standard***
- **DSpace Handbook**
- **Curación** de metadatos con **OpenRefine**

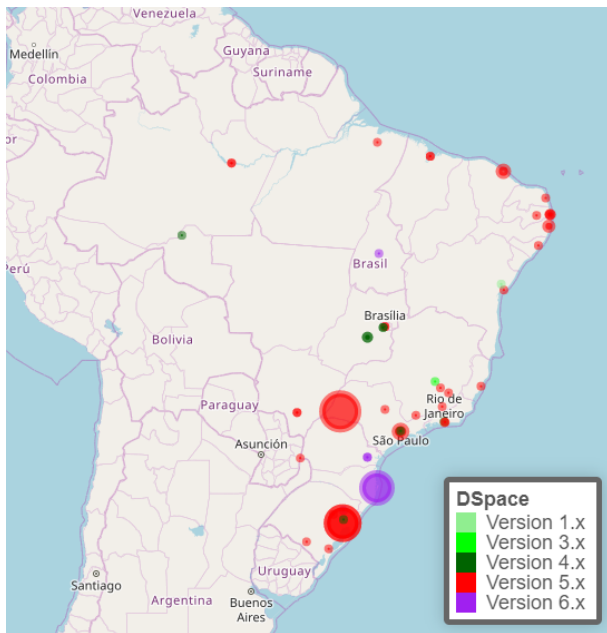
# 135 Instalações do DSpace (90%)



1.x 3.x 4.x 5.x 6.x

meta-chart.com

Version	#	porcentaje
1.x	4	3%
3.x	4	3%
4.x	47	35%
5.x	50	36%
6.x	25	19%



# Sistema de Publicación Electrónica de Tesis y Disertaciones

<https://github.com/ibict-br/TEDE>

English Español Português

Universidade Tuiuti do Paraná  
Biblioteca Digital de Teses e Dissertações

TEDE  
Sistema de Publicação Eletrônica de Teses e Dissertações

Buscar no TEDE

Assuntos com maior número de registros

educação infantil  
qualidade de vida  
cinema  
pedagogical practices  
higher education  
communication  
community social psychology  
educação do campo  
prática pedagógica  
nursing  
linguagem  
public policy  
quality of life  
dysphagia  
políticas públicas  
fonoaudiologia  
audição  
envelhecimento  
trabalho  
educação  
education  
práticas pedagógicas  
práticas pedagógicas  
psicologia social comunitária  
hearing  
deafness

# DSpace Standard (version 6.x)

<https://github.com/ibict-br2/repositorio-padrao>



# RIDI

Repositório Institucional do Instituto Brasileiro  
de Informação em Ciência e Tecnologia

Política do RIDI

Guia de depósito

Buscar no repositório



Autoria



Assunto



Áreas de conhecimento



Tipo de documento



Data de publicação



Tipo de acesso



# config/input-forms.xml

```

<field>
  <dc-schema>dc</dc-schema>
  <dc-element>creator</dc-element>
  <dc-qualifier></dc-qualifier>
  <repeatable>>false</repeatable>
  <label>Autor:</label>
  <type-bind>Tese, Dissertação, Trabalho de Conclusão de Curso</type-bind>
  <input-type>name</input-type>
  <hint>&lt;div class="alert alert-info"&gt;
    Descrição do &lt;strong&gt;AUTOR&lt;/strong&gt;
    &lt;/div&gt;
    &lt;font color="red"&gt;CAMPO OBRIGATÓRIO: &lt;/font&gt; Informe o nome do autor, conforme o formato de citaçã
  <required>&lt;font color="red"&gt;&lt;strong&gt;CAMPO OBRIGATÓRIO:&lt;/strong&gt; Informe o nome do AUTOR, conforme o
    </required>
</field>

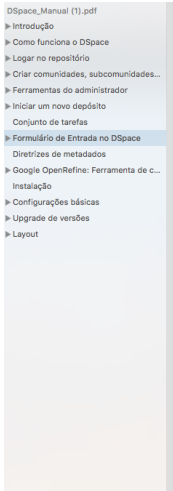
  <field>
    <dc-schema>dc</dc-schema>
    <dc-element>creator</dc-element>
    <dc-qualifier>Lattes</dc-qualifier>
    <repeatable>>false</repeatable>
    <label>&lt;a href="http://buscatextual.cnpq.br/buscatextual/busca.do?metodo=apresentar" target="_blank" &gt;URL do Cur
      <type-bind>Tese, Dissertação, Trabalho de Conclusão de Curso</type-bind>
    <input-type>onebox</input-type>
    <hint>&lt;font color="red"&gt;CAMPO OBRIGATÓRIO: &lt;/font&gt; Informe a URL do currículo lattes do autor.</hint>
    <required>&lt;font color="red"&gt;&lt;strong&gt;CAMPO OBRIGATÓRIO:&lt;/strong&gt; Informe a URL do currículo lattes do
      </required>
  </field>

```



# DSpace Handbook

<https://github.com/ibict-br/DSpace-Manual>



criada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e em outras versões dos formulários de entradas que possam ser utilizadas.

## 8.1 Formulário de Entrada

O formulário de entrada é utilizado para determinar a estrutura dos metadados que serão usados para descrever os documentos durante o processo de submissão.

Em linguagem mais simples e no contexto dos repositórios, os metadados são os campos que serão preenchidos para descrever o objeto digital que está sendo submetido

É um arquivo em XML (eXtensible Markup Language) composto por campos (indicados pela tag `<field>`) que podem ser adequados aos tipos de documentos contidos no repositório, ou seja, você pode configurar o formulário para que este contenha os metadados necessários de acordo com o tipo de documento a ser depositado

Essa configuração **não** pode ser feita através da interface gráfica, sendo necessário o auxílio de equipe com conhecimento técnico do DSpace para que seja executada

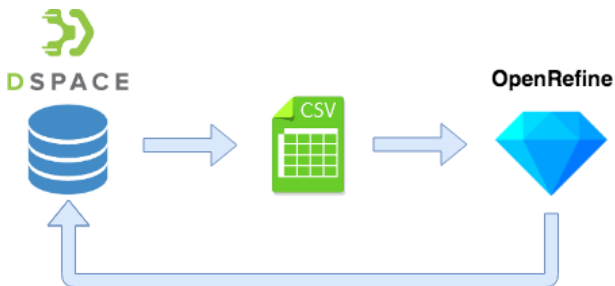
O endereço do arquivo que permite a edição do formulário é o `[dspacebase]/config/input-forms.xml`.

O DSpace usa por padrão o esquema de metadados Dublin Core. Esse formato de metadados dispõe de uma quantidade básica de descritores (Dublin Core Simplificado), mas que pode ser ampliada com qualificadores conforme a necessidade (Dublin Core Qualificado).

Para registrar novos metadados no DSpace é necessário inseri-los também via interface, conforme mostrado em tópicos anteriores.

Veja o exemplo abaixo e as explicações a seguir:

# Curación por medio del OpenRefine



# Descargar OpenRefine

<https://openrefine.org/download.html>



OpenRefine

A free, open source,  
powerful tool for working  
with messy data



Home

Community

Documentation

**Download**

Contact Us

Blog

## Download

You will find on this page a list of OpenRefine distributions and extensions available for download. Are we missing something? Want to fix a typo? You can submit changes (pull request) [from here](#).

## Official Distribution

Read the [installation instructions](#)

You can also download all official releases and source from our [GitHub releases page](#)

## OpenRefine 3.3 beta

The first beta release of OpenRefine 3.3. Please **BACKUP** your workspace directory before installing and report any problems that you encounter.

The first beta release of 3.3 was released on October 21, 2019. A change log is provided on [the release page](#).

- **Windows kit**, Download, unzip, and double-click on *openrefine.exe*. If you're having issues with the above, try double-clicking on *refine.bat* instead.
- **Mac kit**, Download, open, drag icon into the Applications folder and double click on it.
- **Linux kit**, Download, extract, then type `./refine` to start.

# Ejemplos

- 1 Curar URLs
- 2 Migrar elementos entre colecciones
- 3 Normalización de relleno de campos
- 4 *Cluster and Edit*
- 5 Eliminar metadatos